Proc. of the 2017 IEEE International Conference on Signal and Image Processing Applications (IEEE ICSIPA 2017), Malaysia, September 12-14, 2017

Classification of SD-OCT images using a Deep learning approach

Muhammad Awais[1,2*], Henning Müller[3], Tong B. Tang[1,2] and Fabrice Meriaudeau[1,2,*]

[1]Centre for Intelligent Signals and Imaging Research, Universiti Teknologi PETRONAS Malaysia,
[2]Department of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS Malaysia,
[3] University of Applied Sciences Western Switzerland, Sierre (HES-SO) Rue du TechnoPôle 3, 3960 Sierre, Switzerland
mawais069@gmail.com, fabrice.meriaudeau@utp.edu.my

***Abstract*— Diabetic Macular Edema (DME) is one of the many eye diseases that is commonly found in diabetic patients. If it is left untreated it may cause vision loss. This paper focuses on classification of abnormal and normal OCT (Optical Coherence Tomography) image volumes using a pre-trained CNN (Convolutional Neural Network). Using VGG16 (Visual Geometry Group), features are extracted at different layers of the network, e.g. before fully connected layer and after each fully connected layer. On the basis of these features classification was performed using different classifiers and results are higher than recently published work on the same dataset with an accuracy of 87.5%, with sensitivity and specificity being 93.5% and 81% respectively.**

Keywords: Diabetic Macular Edema (DME), Deep learning, Feature Matrices, Visual Graphic Geometry (VGG)

I. INTRODUCTION

Diabetic Macular Edema (DME) is a type of eye disease due to the damage of blood vessels in the retina. When left untreated, DME causes the build-up of liquid in the macula further leading to a swollen area on the retinal layer and as a consequence to irreversible eye blindness. A recent review, based on SD-OCT (Spectral Domain-Optical Coherence Tomography) images [2], conducted by Trichonas and Kaiser [1] highlighted five different patterns of structural changes in DME: sponge-like retinal swelling which is also known as Diffuse Retinal Thickening (DRT), Cystoid Macular Edema (CMD) and Serious Retinal Detachment (SRD), Posterior Hyaloidal Traction (PHT) without Tractional Retinal Detachment (TRD) and PHT with TRD.

Automated diagnosis applied to OCT imaging is still at an early stage, as only academic work has been published and no commercial products are available, yet. Most of the pioneer work on OCT image analysis has focused on the problem of retinal layer segmentation [3, 4] or specific lesion (e.g., cysts) segmentation, as explained in [5, 6]. More recently, SD-OCT databases with their corresponding ground-truth were provided for benchmarking; for instance, a challenge (OPTIMA) was organized as a satellite event of the MICCAI (Medical Image Computing and Computer Assistend Intervention) 2015 conference. Regarding Computer Aided Diagnosis, only little work has been published recently [12, 13, 14, 15]. Some of these articles were based on a set of "hand-crafted" features combining low level and high level features. Dimensionality reduction was done either through Principal Component Analysis or using a bag of words. On a data set of 32 volumes [13, 14, 15], distributed evenly between normal and abnormal cases, the best results obtained were a sensitivity (SE) and specificity (SP) of 87.5% and 87.5%, respectively.

In recent years, deep learning has witnessed significant advances compared with other machine learning techniques. In the field of medical imaging deep learning is one of the most important research areas. Much research has been done on CT (Computed Tomography), MRI (Magnetic Resonance Imaging), PET (Positon Emission Tomography), and X-ray images using deep learning and results are often outperforming other learning algorithms with applications to dermatology [2], prostate cancer classification [8, 9], image registration [10, 11], lung cancer detection [17] and many other applications. Most of these articles are either based on stacked Auto-Encoders or Convolutional Neural Networks. The latter approach uses most commonly either fine-tuning or transfer learning because the training data set are often not large enough to train a deep network from scratch.

This paper focuses on classification of normal and abnormal OCT images using one of the state of the art deep learning networks: VGG-16. Using VGG-16, features are extracted from the SD-OCT images at three locations in the network (after the first, 2nd and 3rd fully connected layers). The images provided to the network are either, raw, denoised, cropped or a combination of these preprocessing steps.

This paper is organized as follows: Section II briefly presents the dataset as well as the deep network used in this study. Section III presents the various experiments and the obtained results. Section IV finally highlights the main results obtained.

II. DATASET AND NETWORK

2.1 Dataset

The dataset used in the proposed algorithm obtained ethics approval and was acquired by the Singapore Eye Research Institute (SERI), using a CIRRUS TM (Carl Zeiss Meditec, Inc., Dublin, CA) SD-OCT device [13, 14, 15]. The dataset consists of 32 OCT volumes (16 DME and 16 normal cases). Each volume contains 128 B-scans with a resolution of 1,024 px × 512 px. All SD-OCT volumes were read and assessed by trained graders and identified as normal or DME based on evaluation of retinal thickening, hard exudates, intraretinal

cystoid space formation, and subretinal fluid within the DME sub-set.

2.2 VGG Network and feature extraction

K. Simonyan and A. Zisserman [18] proposed a very deep convolutional network for large-scale image recognition (VGG). They designed several VGG models, such as VGG19, VGG16, VGG13, VGG11 as shown in Fig. 1. The best obtained 92.7% top-5 test accuracy on the ImageNet dataset, that comprises over 14 million images belonging to 1000 classes.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
Input (224 x 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure.1: VGG ConvNet configuration.

In this paper we used the macro architecture of VGG16 as can be seen in Fig. 2.

The input to our VGG ConvNets is of fixed-size of 224 × 224 pixels as RGB image, where each color channel receives a copy of the same BScan of our SD-OCT volumes. Then, the input image is passed through a stack of convolutional (conv.) layers. The convolution stride is fixed to 1 pixel; the spatial padding of conv. layer input is such that the spatial resolution is preserved after convolution, i.e. the padding is 1 pixel for 3 × 3 convolution layers. Spatial pooling is carried out by five max-pooling layers that follow some of the conv. layers (not all the conv. layers are followed by max-pooling). Max-pooling is performed over a 2 × 2-pixel window, with stride 2. A stack of convolutional layers (that have a varying depth in different architectures) is followed by three fully-connected (FC) layers: the first two have 4096 channels each, the third performs 1000-way ILSVRC (ImageNet Large Scale Visual Recognition Competition) classification and thus contains 1000 channels (one for each class). The final layer is the soft-max layer. The configuration of the fully connected layers is the same in all networks. All hidden layers are equipped with the rectification (ReLU) non-linearity activation function.

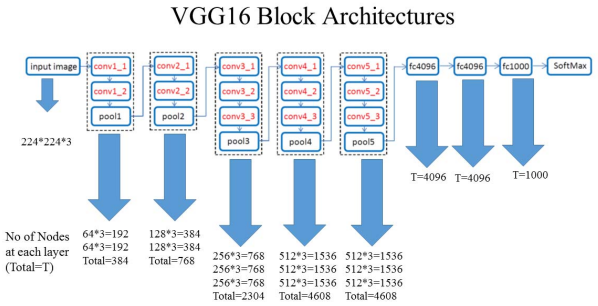

Figure.2 Microarchitecture of VGG16.

III. Experiments and Results

3.1 Classification and Evaluation

At this stage a kNN classifier (with K=1 and 3) and a Random Forest classifier (100 trees) were tested, using the feature vector provided by the VGG network with size ranging from 4096 to 1000 depending at which level of the FCC the classifier is connected. The evaluation is performed using a leave-one-patient-out patient. As each BScan is evaluated, a majority rule is employed to classify the entire volume.

For evaluation, all results are expressed in terms of Sensitivity (SE) and Specificity (SP).

- Sensitivity (SE) – The ability of a test to correctly identify those with DME disease.
- Specificity (SP) – The ability of a test to correctly identify those without DME disease.

3.2 Experiments

Four experiments were conducted on the dataset with different preprocessing. There are a total of 16 volumes each for DME and normal patients.

Experiment #1

Experiment #1 is carried out on raw datasets with no noise removal and without image cropping (A), where the layers are detected using the algorithm presented in [12]. Fig.3 shows examples of input images for DME and normal patients. Table 1 shows the obtained results for the classifiers and levels in the FCC.

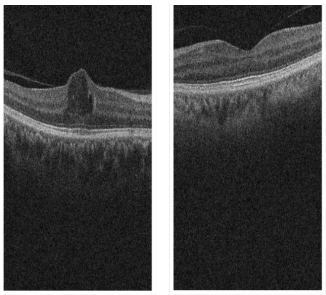

Fig. 3 Raw dataset. (a) DME patient (b) normal patient

Table.1 Classication results without noise removal and without image cropping (A) *after the* 1st Fully Connected Layer (1-FCL), 2nd Fully Connected Layer (2-FCL) and the 3rd Fully Connected Layer (3-FCL)

	1-FCL			2-FCL			3-FCL		
A	Ac	Se	Sp	Ac	Se	Sp	Ac	Se	Sp
	K-NN(k=1)								
	87 %	93 %	81 %	93%	87 %	100 %	45 %	35 %	64 %
	K-NN(k=3)								
	65 %	81 %	43 %	87.5 %	93 %	81%	84 %	81 %	87 %
	Decision Tree								
	87 %	93 %	81 %	75%	93 %	65%	84 %	93 %	75

Experiment #2

Experiment #2 is carried out on the dataset with noise removal but without image cropping. Fig. 4 shows an example of an input image for DME and a normal patient. Table 2 shows the obtained results for the classifiers and levels in the FCC.

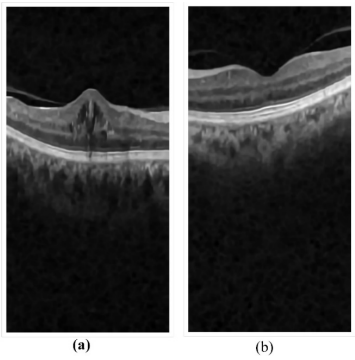

Fig.4 Dataset with noise removal but without image cropping. (a) DME patient (b) normal patient.

Table.2 Classication results with noise removal but without image cropping (B) *after* Ist Fully Connected Layer (1-FCL), 2nd Fully Connected Layer(2-FCL) and 3rd Fully Connected Layer(3-FCL)

	1-FCL			2-FCL			3-FCL		
B	Ac	Se	Sp	Ac	Se	Sp	Ac	Se	Sp
	K-NN(k=1)								
	68.5%	37.5 %	100 %	87.5%	100 %	75%	87.5	75 %	100
	K-NN(k=3)								
	65 %	31.5 %	100	90.6%	100 %	81.25%	87.5	81.5%	100 %
	Decision Tree								
	71 %	81.25%	43 %	90.6%	93.5%	87.5 %	90.6	93.5	87.5

Experiment #3

Experiment #3 is carried out on raw datasets with no noise removal but with image cropping. Fig.5 shows the example of input image for DME and normal patient. Table.3 shows the obtained results for the different classifiers and different levels in the FCC.

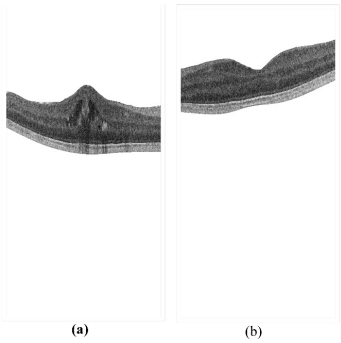

Fig. 5 Dataset with no noise removal but with image cropping. (a) DME patient (b) normal patient.

Table.3 Classication with no noise removal but with image cropping (C) *after the* 1st Fully Connected Layer (1-FCL), the 2nd Fully Connected Layer (2-FCL) and the 3rd Fully Connected Layer (3-FCL)

	1-FCL			2-FCL			3-FCL		
C	Ac	Se	Sp	Ac	Se	Sp	Ac	Se	Sp
	K-NN(k=1)								
	71 %	43% *	100 %	84%	93.5 %	75 %	84%	100 %	86 %
	K-NN(k=3)								
	76 %	31%	100 %	87.5 %	93.5	81 %	81%	100 %	62 %
	Decision Tree								
	75 %	93.5 %	65 %	87%	100 %	75 %	84.5 %	93.50	75 %

Experiment #4

Experiment #4 is carried out on the dataset with noise removal and image cropping. Fig. 6 shows an example of an input image for DME and a normal patient. It can be seen that all the harsh edges are smoothed and a clearer image of the retinal layer is seen. Moreover, the irrelevant parts are also excluded from feature extraction. Table 4 shows the results obtained for the classifiers and levels in the FCC.

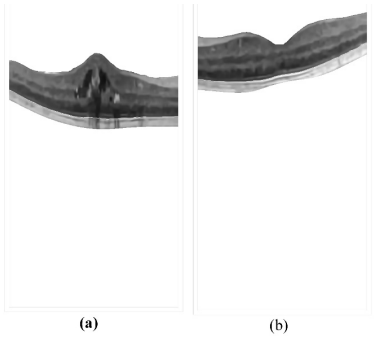

Fig. 6 Dataset with noise removal and image cropping. (a) DME patient (b) normal Patient

Table.4 Classication with noise removal and image cropping (D) *after* 1st Fully Connected Layer (1-FCL), the 2nd Fully Connected Layer (2-FCL) and the 3rd Fully Connected Layer (3-FCL)

	1-FCL			2-FCL			3-FCL		
D	Ac	Se	Sp	Ac	Se	Sp	Ac	Se	Sp
	K-NN(k=1)								
	75 %	53.3 %	100 %	87 %	87%	87 %	87	100	75
	K-NN(k=3)								
	53 %	6%	100 %	84 %	81.5 %	87 %	87	100	75
	Decision Tree								
	65 %	81	43%	84 %	81.50	87 %	84 %	93.5	75 %

IV. Conclusion

The use of of OCT provides high resolution retinal images for DME detection and the adaptation of deep learning has proven to improve image classification with high performance, such as an accuracy of 90%. Deep learning application on the task of DME detection using VGG16 increased the SE performance by more than 20% compared to previous researche. This opens a new, simple and effective method for early DME detection to aid ophthalmologists in diagnosis.

For future work, dimension reduction approaches such as PCA or BoW will be investigated, as well as a combination of several deep learning architectures using a majority voting approach. Moreover, fine-tuning of the network will be investigated.

V. References

[1] G. Trichonas and P. K. Kaiser, "Optical coherence tomography imaging of macular oedema," British Journal of Ophthalmology, vol. 98, pp. ii24-ii29, 2014.

[2] D.VijayJoshi, "Optical Coherence Tomography" Available online: https://www.slideshare.net/ vijayjoshi311/optical-coherence-tomography-26825183.Accessed Date:05-11-2017

[3] J. G. Fujimoto and E. A. Swanson, "The Development, Commercialization, and Impact of Optical Coherence Tomography," Inves. Ophthalmol. Visual Sci. (2016).

[4] E. Swanson, "Beyond Better Clinical Care: OCT's Economic Impact," BioOptics World (2016).

[5] E. A. Swanson and D. Huang, "Ophthalmic OCT Reaches $1 Billion per Year: But Reimbursement Clampdown Clouds Future Innovation," Retinal Physician 45, 58–59 (2011).

[6] E. A. Swanson, "Estimates of Ophthalmic OCT Market Size and the Dramatic Reduction in Reimbursement Payments," http://www.octnews.org/articles/4176266/estimates-of-ophthalmic-oct-market-size-and-the-dr/, (2012).

[7] A. E. Fung, G. A. Lalwani, P. J. Rosenfeld, S. R. Dubovy, S. Michels, W. J. Feuer, C. A. Puliafito, J. L. Davis, H. W. Flynn, Jr., and M. Esquiabro, "An optical coherence tomography-guided, variable dosing regimen with intravitreal ranibizumab (Lucentis) for neovascular age-related macular degeneration," Am. J. Ophthalmol. 143(4), 566–583 (2007).

[8] Guo Y, Gao Y, Shen D. 2016. Deformable MR prostate segmentation via deep feature learning and sparse patch matching. IEEE Trans. Med. Imaging 35:1077–89

[9] Liao S, Gao Y, Shi Y, Yousuf A, Karademir I, et al. 2013. Automatic prostate MR image segmentation with sparse label propagation and domain-specific manifold regularization. Inf. Proc. Med. Imaging 23:511–23

[10] Wu G, Kim M, Wang Q, Munsell BC, Shen D. 2016. Scalable high-performance image registration framework by unsupervised deep feature representations learning. IEEE Trans. Biomed. Eng. 63:1505–16

[11] Wu G, Kim M, Wang Q, Gao Y, Liao S, Shen D. 2013. Unsupervised deep feature learning for deformable registration of MR brain images. In Proceedings of the 2013 Medical Image Computing and Computer-Assisted Intervention Conference, pp. 649–56. Berlin: Springer.

[12] S. P. K. Karri, D. Chakraborty, and J. Chatterjee, "Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration," Biomedical Optics Express, vol. 8, pp. 579-592, 2017/02/01 2017.

[13] D Sidibé, S Sankar, G Lemaître, M Rastgoo, J Massich, CY Cheung and F. Meriaudeau, "An anomaly detection approach for the identification of DME patients using spectral domain optical coherence tomography images", Computer Methods and Programs in Biomedicine 139, 109-117.

[14] G. Lemaıtre, M. Rastgoo, J. Massich, C. Y Cheung, T. Y Wong, E. Lamoureux, D. Milea, F. Mériaudeau, D. Sidibé, "Classification of SD-OCT Volumes using Local Binary Patterns: Experimental Validation for DME Detection", Journal of Ophthalmology, Mai 2016.

[15] K. Alsaih, G. Lemaıtre, M. Rastgoo, J. Massich, D. Sidibé F. Mériaudeau, , "Machine Learning Techniques for Diabetic Macular Edema (DME) Classification on SD-OCT images", Biogineering onLine, Accepted mai 2017.

[16] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks", nature letters, vol 542, pp. 115-118, 2017.

[17] S. Wenqing, Z. Bin; Q. Wei, "Computer aided lung cancer diagnosis with deep learning algorithms", Proceedings of the SPIE, Volume 9785, id. 97850Z 8 pp. (2016).

[18] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition" , ILSVRC-2014 .